

A Study of the Effectiveness of Milestones™

Pierce Middle School, Tampa, Florida

2008-2009



This study was conducted by SEG Research, an independent educational research firm located in New Hope, Pennsylvania. SEG Research provides research, evaluation, and assessment services to educational publishers, educational technology providers, assessment service providers and government agencies. SEG has been meeting the research and assessment needs of organizations since 1979. This research was supported by a grant from Heinle.

A Study of the Effectiveness of Milestones

Pierce Middle School, Tampa, Florida

Executive Summary

Background and Purpose

During the 2008-2009 school year, SEG Research conducted a study of the effectiveness of *Milestones*, an instructional program designed to provide intervention for English Language Learners (ELL) and Struggling Readers. The study examines the growth in Reading and Language skills for students enrolled in a class of 6th grade students in a Developmental Language Arts through ESOL Program at Pierce Middle School in Tampa, Florida. These students received instruction using *Milestones* (Treatment) and their Reading and Language proficiency growth was measured at the end of the 2007-2008 school year and then again after receiving instruction at the end of the 2008-2009 school year using a post hoc, pre-post, treatment-only research design. Student Reading and Language skills were measured using the CELLA test required in Florida.

This study investigated the following questions:

1. Do ELL students show significant gains in Reading and Language skills after receiving instruction using *Milestones*?
2. Do boys and girls receiving instruction using *Milestones* differ in their level of Reading and Language skills growth?
3. Do students of different ethnicities receiving instruction using *Milestones* differ in their level of Reading and Language skills growth?
4. Do students participating and not participating in free or reduced lunch receiving instruction using *Milestones* differ in their level of Reading and Language skills growth?

Study Design

Students receiving instruction using *Milestones* participated in an effectiveness study using a post hoc, pre-post, Treatment- Group only design. Student growth in Language and Reading skills was

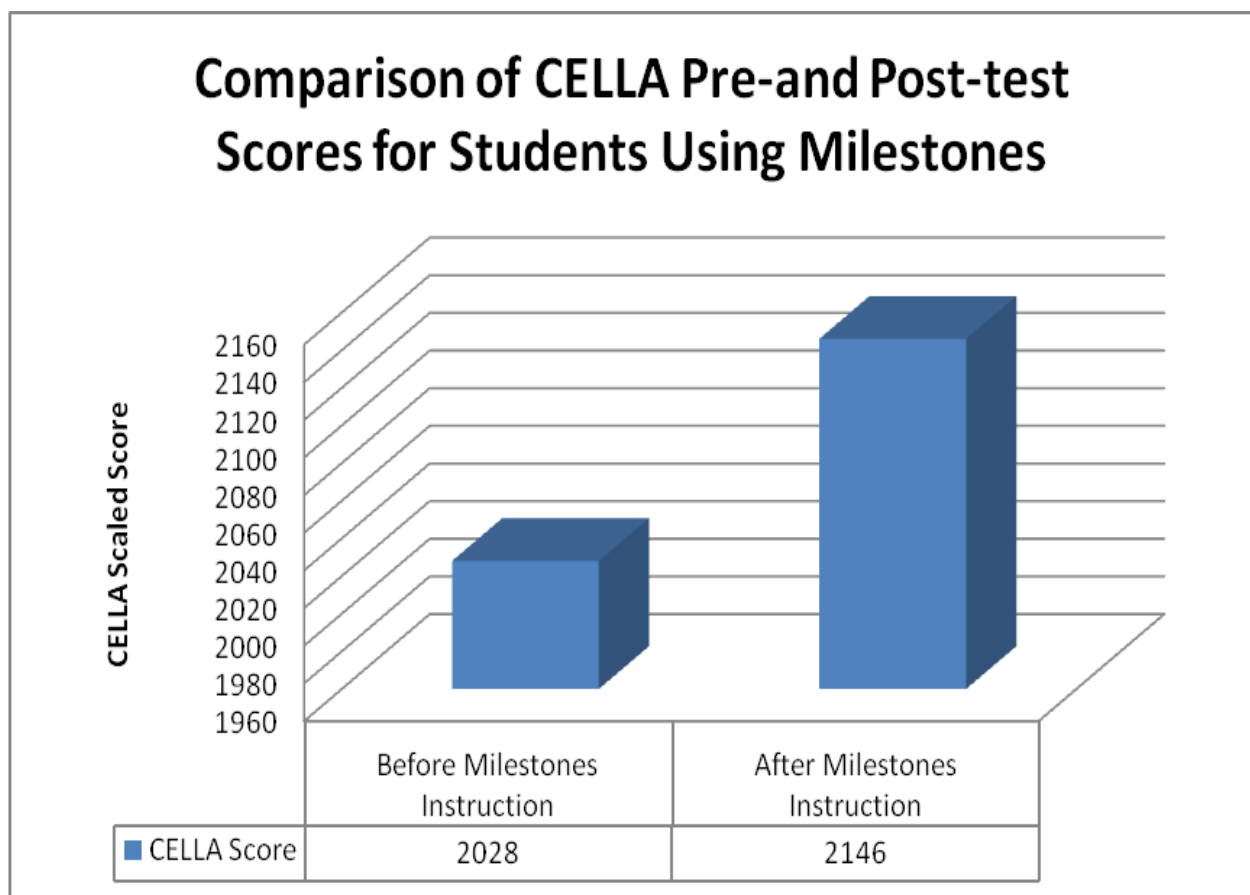
A Study of the Effectiveness of Milestones

measured by comparing their proficiency at the beginning of instruction and again after receiving instruction. Students were administered CELLA as a pre-test at the end of the 2007-2008 school year and again at the end of the 2008-2009 school year after receiving instruction. The pre- and post-test results were then compared statistically.

Findings

Pre-Post Growth

Students provided with instruction using *Milestones* showed substantial growth from pre- to post-test. During the course of the study, students showed a statistically significant increase in their CELLA scores. This increase is illustrated below.



To better understand the magnitude of growth for students in *Milestones* classrooms we looked at the “effect size”, a common metric that can be used to evaluate the amount of growth across studies, when different measures are used; the effect size for growth in CELLA scores was $+0.81$. This indicates that students provided with instruction using *Milestones* showed substantial growth in Reading and Language skills as indicated by CELLA scores. While effect size is a relative measure and there is no absolute criterion for what constitutes a large effect, for a student at the 50th percentile, this would represent an increase to the 79th percentile from pre-to posttest.

Growth in Proficiency Levels

Based on the CELLA test scores, students taking the test are classified into one of five proficiency levels based on the California TESOL classifications: Beginning, Early Intermediate, Intermediate, Early Advanced and Advanced. More than half (52%) of the students gained one or more proficiency levels. Of those, about one fifth (18%) gained more than one level moving from A to C. the remaining third (34%) moved one level from A to B or B to C.

Gender Analysis

We examined whether there were any differences in growth between male and female students,

Milestones was equally effective for both male and female students. There were no significant differences in CELLA post test scores between male and female students ($F=.236$; $df(2,22)$; $p>.05$). Female students had an adjusted mean post test score of 2140.30 and male students had an adjusted mean post test score of 2153.34. (All students were participating in the free and reduced lunch program and ethnicity was not reported, so subgroup analyses were not conducted for these background variables.)

SUMMARY

The students using *Milestones* showed significant gains in Reading and Language skills over the course of the 2008-2009 school year. Students showed substantial increases in CELLA scores from the pre- to post-test and more than half of the students increased by one or more proficiency levels. There were no significant differences in CELLA post-test scores between boys and girls, indicating that *Milestones* was equally effective for both genders.

A Study of the Effectiveness of Milestones

Pierce Middle School, Tampa, Florida

Milestones Program Overview

Milestones is an instructional program designed to provide intervention for English Language Learners (ELL) and Struggling Readers. Using a unique embedded assessment plan along with a balanced blend of literature and content Readings, *Milestones* ensures that students are mastering skills and standards before being introduced to new skills and standards. This consistent, research-based approach supports student Language acquisition and content knowledge crucial for academic success.

Effectiveness Study Goals and Overview

This report describes a study conducted during the 2008-2009 school year to evaluate the effectiveness of *Milestones*. The study examines the growth in Reading and Language skills for students enrolled in a class of 6th grade students in a Developmental Language Arts through ESOL Program at Pierce Middle School in Tampa, Florida. These students received instruction using *Milestones* (Treatment) and their Reading and Language proficiency growth was measured before and after receiving instruction using a post hoc, pre-post, treatment-only research design. We compared student growth by examining student Language skills over the course of a year using the Florida required CELLA measure of ESL language skills.

Research Questions

This study investigated the following questions:

1. Do ELL students show significant gains in Reading and Language skills after receiving instruction using *Milestones*?
2. Do boys and girls receiving instruction using *Milestones* differ in their level of Reading and Language skills growth?
3. Do students of different ethnicities receiving instruction using *Milestones* differ in their level

A Study of the Effectiveness of Milestones

of Reading and Language skills growth?

4. Do students participating and not participating in free or reduced lunch receiving instruction using *Milestones* differ in their level of Reading and Language skills growth?

Student Sample

Thirty-seven students enrolled in the Developmental Language Arts through ESOL Program using *Milestones* at Pierce Middle School in Fairfield, California participated in the study. Half of the participating students were Male (18) and half were female (19). All (37) students in the study were receiving free or reduced lunch. The ethnicity of the students was not reported.

In some cases, as a result of students entering or exiting the district during the school year, complete background information for a student may not be available or a student did not take one of the tests included in the analyses. Where data was missing, the student's results were eliminated from those analyses.

Description of the Pre-test and Post-test

Student Reading and Language skills were measured using the CELLA Test. The CELLA is administered to ELL students annually as a basis for measuring students' English language skills as a basis for monitoring progress and determining readiness for exit from an ELL program..

The CELLA was used for both pre- and post-testing in this study. The CELLA is administered toward the end of each school year. The 2007-2008 CELLA scores served as the pre-test measure for this study and the 2008-2009 CELLA scores served as the post test measure for the study. The study therefore measures the growth occurring between the end of the 2007-2008 school year and the end of the 2008-2009 school year. The Florida Department of Education reports the reliability of the CELLA ranges from .82 to .89, depending on grade level.

Study Design

The goal of this effectiveness study was to evaluate student growth in Language and Reading skills after receiving instruction using *Milestones*. The study employed a post hoc, pre-post, treatment-group only design. Student growth in Language and Reading skills was measured by comparing their proficiency at the beginning of instruction and again after receiving instruction. Students were administered the CELLA as a pre-test at the end of the 2007-2008 school year and as a post test at the end of the 2008-2009 school year. The pre- and post-test results were then compared statistically.

Data Collection

At the outset of the study, we collected background information about the participating students in order to characterize the sample, and to facilitate comparisons across population subgroups. This information included:

- Student ethnicity (not reported)
- Student gender
- Student free or reduced lunch participation

Thirty-seven students enrolled in the Developmental ELL Program using *Milestones* at Pierce Middle School in Tampa, Florida participated in the study. Working with teachers and administrators at Pierce Middle School, we obtained demographic and background information as well as CELLA scores for the students participating in the study. Both the results for 2007-2008 and 2008-2009 CELLA results were obtained. Student ethnicity was not reported. In some cases, as a result of students entering or exiting the district during the school year, student test results may not have been available for one of the years of interest; in these cases, the student's results were eliminated from those analyses.

Findings

Measuring Growth

The overall growth in Reading and Language skills following instruction using *Milestones* was determined using a t-test. This statistic was used to compare the mean (average) Reading and Language skill level before instruction to the average scores after instruction as measured by the CELLA.

After evaluating the overall growth of students, we examined potential differences in growth levels between girls and boys. These comparisons were made using a statistical procedure known as analysis of covariance (ANCOVA). This approach provides an accurate way to compare growth over time controlling for any potential differences in student skills between gender, ethnic and free or reduced lunch groups that may have been present at the beginning of the study. Any differences in skill levels between boys and girls, different ethnic groups and those eligible and ineligible for reduced school lunch that may have existed at the beginning of the study were controlled to ensure that any differences in subsequent growth were the result of *Milestones* use and not merely the result of differences between those groups that existed at the start of the study. While no procedure can completely eliminate differences that may exist at the outset of a study, ANCOVA is widely recognized as an effective way to control for differences.

Only students for whom matched pre-test and post-test results were available were included in the analysis. The analysis looked only at those students who had taken CELLA at the beginning of the study (pre-test) and those who had taken the CELLA at the end of the study (post-test). Students who left the class during this period or who joined the class during this period were not included in the growth comparisons.

Pre-Post Growth

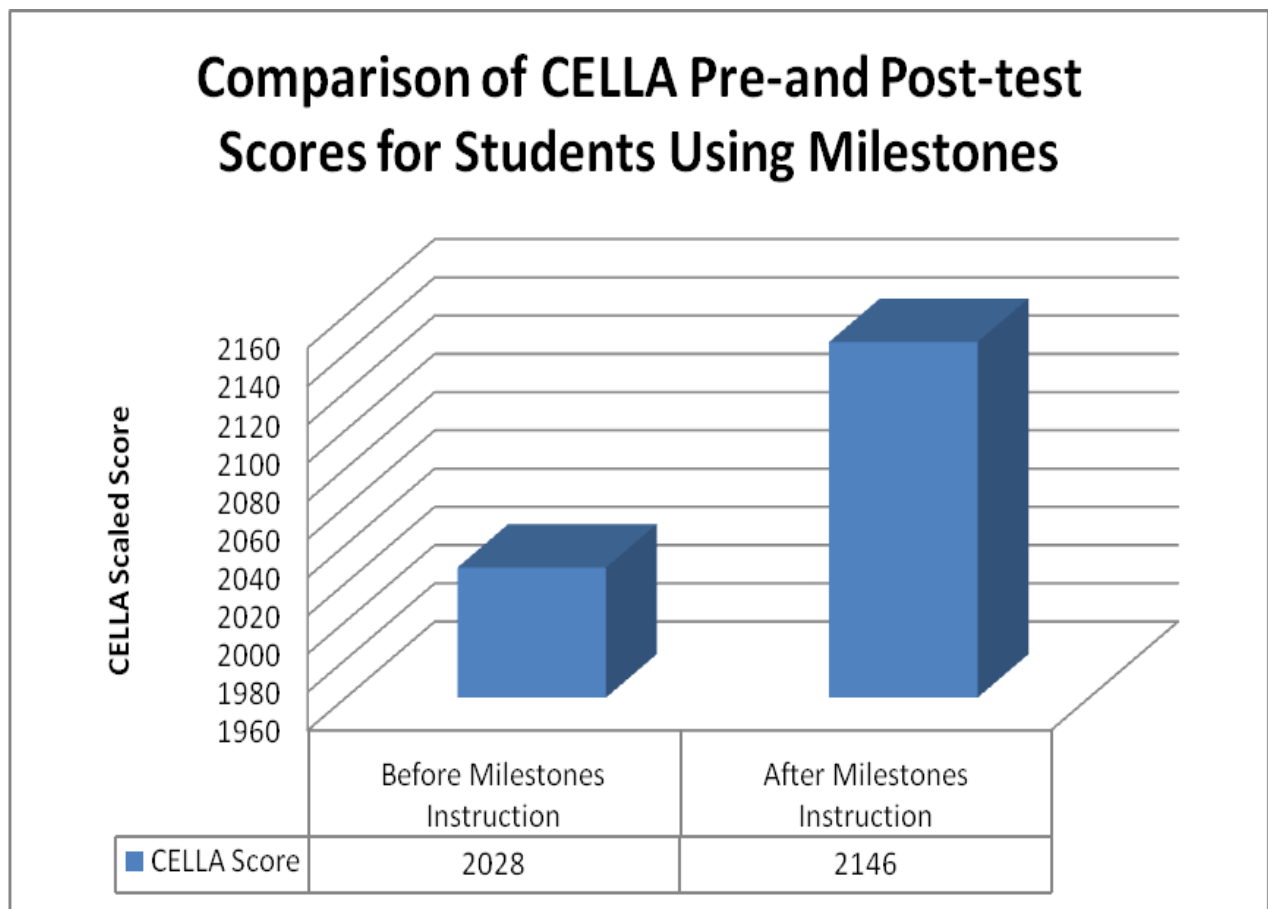
Students provided with instruction using *Milestones* showed substantial growth from pre- to post-test on the CELLA. Twenty-two students took both the pre and post test. The mean CELLA score in 2007-2008 was 2027.64 and the mean CELLA score in 2008-2009 was 2146.23. This represents a

A Study of the Effectiveness of Milestones

gain of approximately 118.59 points. The CELLA scores from pre- to post-test were significantly different ($t= 7.51$; $df(21)$; $p<.001$). This means that there was less than a one in a thousand chance that the differences observed were due to chance.

To better understand the magnitude of growth for students in *Milestones* classrooms we looked at the “effect size”, a common metric that can be used to evaluate the amount of growth across studies, when different measures are used. The effect size (calculated as the difference in the pre and post test means divided by the post test standard deviation) was $+ .81$.

This indicates that students provided with instruction using *Milestones* showed substantial growth in Reading and Language skills as indicated by CELLA.



CELLA Proficiency Levels

It is also interesting to examine student growth in proficiency level classification as well as the test scores. Based on the CELLA test scores, students taking the test are classified into one of five proficiency levels based on the TESOL classifications: Beginning, Early Intermediate, Intermediate, Early Advanced and Advanced. All of the students in this study were in the range of Beginning to Intermediate.

Based on the test score growth during the course of the study, students typically rose in proficiency level during the course of the study. More than half (52%) of the students gained one or more proficiency levels. Of those, about one fifth (18%) gained more than one level moving from A to C. the remaining third (34%) moved one level from A to B or B to C. the majority of the students showed gains in proficiency level, even when students showed significant gains in test scores, it may not have been enough to move them to the next proficiency level category.

Gender Analysis

We examined whether there were any differences in growth among male and female students. All students were receiving free and reduced lunch and ethnicity was not reported, so analyses were conducted for these background variables.

We compared the growth in Reading and Language skills as measured by the CELLA scores using analysis of covariance (ANCOVA). We examined the differences in CELLA post test scores (dependent measure) between subgroups (independent measure) controlling for the initial skill levels of the students reflected in the pre-test scores (covariate). The CELLA pre-test scores were used as the covariate to place students in each population subgroup on the same baseline.

Milestones was equally effective for both male and female students. There were no significant differences in CELLA post test scores between male and female students ($F=.236$; $df(2,22)$; $p>.05$). Female students had an adjusted mean post test score of 2140.30 and male students had an adjusted mean post test score of 2153.34.

A Study of the Effectiveness of Milestones

Table 3: Analysis of Covariance: Differences in CELLA Post-Test Scores Among Gender, Ethnic and Free or Reduced Lunch Eligibility Groups for Students Receiving Instruction Using *Milestones*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	75585.513 ^a	2	37792.757	12.149	.000	.561
Intercept	50096.731	1	50096.731	16.105	.001	.459
CELLA 2007-2008 Scores (Covariate)	65437.416	1	65437.416	21.037	.000	.525
Gender	734.816	1	734.816	.236	.633	.012
Error	59102.350	19	3110.650			
Total	1.015E8	22				
Corrected Total	134687.864	21				

SUMMARY AND DISCUSSION

This report describes a study conducted during the 2008-2009 school year to evaluate the effectiveness of *Milestones*. The study examined the effectiveness of *Milestones*, an instructional program designed to provide intensive intervention for English Language Learners (ELL). The study examined the growth in Reading and Language skills of students enrolled in a class of 6th grade students in a Developmental Language Arts through ESOL Program at Pierce Middle School in Tampa, Florida. These students received instruction using *Milestones* (Treatment) and their Reading and Language proficiency growth was measured at the end of the 2007-2008 and at the end of 2008-2009.

Study Design

The study investigated whether ELL students achieve significant gains in Reading and Language skills after receiving instruction using *Milestones* using a post hoc, pre-post, treatment group only design. Additionally, potential differences among gender, ethnic and free and reduced lunch groups were explored.

Student growth in Language and Reading skills was measured by comparing their proficiency at the beginning of instruction and again after receiving instruction. Student proficiency was measured using the Florida required measure of ELL Reading and Language skills, CELLA. Students were administered the CELLA as a pre-test at the end of the 2007-2008 school year and again at the end of the 2008-2009 school year. The pre- and post- test results were compared statistically.

Thirty-seven students enrolled in the Developmental Language Arts through ESOL Program using *Milestones* at Pierce Middle School in Fairfield, California participated in the study. Working with teachers and administrators at Pierce Middle School, we obtained demographic and background information as well as CELLA and CST scores for the students participating in the study. Both the results for 2007-2008 and 2008-2009 were obtained. In some cases, as a result of students entering

A Study of the Effectiveness of Milestones

or exiting the district during the school year, student test results may not have been available for one of the years of interest; in these cases, the student's results were eliminated from those analyses.

Summary of Results

Pre-Post Growth for *Milestones* users

The students showed significant gains in Reading and Language skills after receiving intervention using *Milestones*. During the course of the study, students increased their CELLA scores by more than 76 points; the effect size for growth in CELLA scores was +.81. This indicates that students provided with instruction using *Milestones* showed substantial growth in Reading and Language skills as indicated by CELLA.

Growth in Proficiency Levels

Based on the CELLA test scores, students taking the test are classified into one of five proficiency levels based on the California TESOL classifications: Beginning, Early Intermediate, Intermediate, Early Advanced and Advanced. Nearly two thirds (62%) of the students gained one proficiency level moving from Beginning to Early Intermediate or from Early Intermediate to Intermediate.

Based on the test score growth during the course of the study, students typically rose in proficiency level during the course of the study. More than half (52%) of the students gained one or more proficiency levels. Of those, about one fifth (18%) gained more than one level moving from A to C. the remaining third (34%) moved one level from A to B or B to C.

Gender Comparisons

There were no significant differences in CELLA post test scores between male and female students after controlling for initial skill levels. *Milestones* was equally effective for both male and female students.

Conclusion

Students showed significant growth in Reading and Language skills after receiving instruction using *Milestones*. Moreover, the size of the effect observed (.81) is substantial. While effect size is a relative measure and there is no absolute criterion for what constitutes a large effect, for a student at the 50th percentile, this would represent an increase to the 79th percentile from pre-to posttest.

While these findings provide evidence of the effectiveness of *Milestones*, additional research in more controlled settings would shed further light on the effectiveness of *Milestones* and would further isolate *Milestones* as the cause of the large gains in student performance witnessed in this study. Moreover, further research with larger groups of students would be useful in further validating these findings.

The findings of this study demonstrate that students provided with intervention using *Milestones* can make significant gains in Reading and, Language skills.